Грум-Гржимайло Юрий Владимирович
кандидат экономических наук.,
зав. сектором механизмов финансирования
и форм организации науки РИЭПП.
(495) 916-14-79, info@riep.ru

К ВОПРОСУ О ПРИМЕНЕНИИ ИНФОРМАЦИОННЫХ ПОДХОДОВ К АНАЛИЗУ ТЕНДЕНЦИЙ РАЗВИТИЯ НАУЧНО-ТЕХНИЧЕСКОЙ СФЕРЫ

Ключевые слова: информационные методы, анализ, научно-техническая сфера.

Поводом для написания данной статьи стало наблюдаемое разделение между экспертными, экономико-статистическими и информационными методами анализа тенденций развития научно-технической сферы[1]. Традиционно считается, что сложные вопросы анализа тенденций развития являются предметом экспертных оценок, форсайта и других экспертных технологий, а статистические исследования в данном случае во многом предопределены степенью полноты имеющейся статистической базы. Вопросы об информационных подходах к анализу тенденций развития научно-технической сферы рассматривались эпизодически, от случая к случаю, и в основном сводились к вопросам формирования информационной базы для аналитических исследований и многочисленным призывам использовать имеющийся информационный потенциал.

По мере развития информационной инфраструктуры научных исследований и возможностей поиска информации в слабоструктуризованных документальных массивах начали формироваться различные методы «информационных» измерений, например библиометрические методы оценки развития того или иного научного направления по числу научных публикаций и т. д. Подобные оценки, на наш взгляд, очень важны и помогают дополнить оценки официального статистического наблюдения. В публикуемых статистических справочниках можно выделить несколько уровней отражения научно-технической сферы в статистическом наблюдении:

1) уровень продукции, продуктов и услуг научно-технической сферы;

2) уровень видов экономической деятельности в научно-технической сфере;

3) уровень описания научно-технической деятельности в региональном разрезе;

---

[1] Автор выражает благодарность к.ф.-м.н. В. В. Борисову и к.т.н. С. В. Попову за обсуждение и ценные замечания, сделанные в процессе подготовки данной публикации.

4) уровень описания социума и кадров научно-технической сферы;

5) уровень обобщенных (синтетических) показателей развития научно-технической сферы.

Если показатели статистического наблюдения научно-технической сферы тяготеют к обобщенности, «макроуровню» ее представления, то «информационные» измерения могут быть полезны на «микроуровне» исследования, когда требуется различать конкретные объекты или процессы, которые «теряются» в обобщенной статистической информации. С другой стороны, «информационные» измерения предполагают наличие модельного представления об изучаемом объекте или процессе, которое называют «информационной моделью». История понятия «информационная модель» имеет давнюю историю. Оно всегда было связано с познавательной деятельностью, но с развитием компьютеров приобрело особый смысл. Многие прикладные задачи информационного моделирования были поставлены и изучены в 50–60-х годах прошлого века в связи с попытками смоделировать на компьютерах психическую деятельность человека при решении творческих интеллектуальных задач. Научное знание и модели, которые были получены в ходе решения этих задач, объединены в научном направлении «Исследования по искусственному интеллекту». Концепцию информационной модели (англ. information model) выдвинули В. П. Зинченко и Д. Ю. Панов (1962), термин «информационная модель» был введен ими в 1964 г. [1]. Понятие «информационной модели» рассматривалось и В. М. Глушковым [2]. В самом общем случае информационная модель многими исследователями с небольшими вариациями определяется как записанная в той или иной форме (символьной, текстовой, графической) структура или совокупность знаний об исследуемом объекте. К такому определению близко стандартизованное понятие информационной модели как «формальной модели ограниченного набора фактов, понятий или инструкций, предназначенная для удовлетворения конкретному требованию» [3].

Поясним сказанное. Допустим, что нас интересует развитие нанотехнологий как нового перспективного направления. Это сложный комплекс физических, химических и иных исследований, поэтому получить данные о количестве публикаций по нанотехнологиям и оценить их динамику можно только имея представление о структуре входящих в данный комплекс научных направлений, поскольку оценка по общему термину «нанотехнология» не даст полноты. Одновременно мы хотим получить данные по статистическим наблюдениям ресурсов, выделяемых на «развитие исследований в области нанотехнологий», но имеем лишь общую оценку по научным исследованиям и разработкам в целом. Как быть? Возникает, на наш взгляд, некоторое противоречие – обобщенное статистическое наблюдение не может показать, что же на самом деле происходит «в области нанотехнологий», а информационные измерения потребуют изучения структуры научных направлений в данной области, что само по себе представляет исследовательскую задачу.

Сочетание различных методов анализа всегда было связано с поиском «золотой середины» и выделением для каждого из используемых методов области применения. Не менее важной проблемой является

проблема сопоставимости данных, которая часто решается на основе использования относительных величин. Так, данные о темпах изменения конкретных показателей можно считать сопоставимыми вне зависимости от того, получены они на основе официального статистического наблюдения или информационного измерения. Важнее оказывается наличие или отсутствие объективной связи между показателями, что опять-таки выявляется в результате структурных исследований предмета проводимого анализа. В экономических исследованиях корреляция показателей обычно трактуется как следствия существования связи между объектами, с которыми соотнесены данные показатели. На основе описания объектов делается заключение о правомерности существования связи. А если посмотреть на технологию исследования наоборот? Скажем, на основе имеющихся источников информации с помощью компьютерных технологий создать структурную информационную модель исследуемой предметной области (например, модель предметной области федеральной целевой программы или критической технологии), соотнести элементы модели с конкретными статистическими показателями и на основе уже выявленных связей анализировать их динамику и корреляцию? Сам собой напрашивается вывод о целесообразности такого подхода для мониторинга развития отдельных приоритетных направлений в научно-технической сфере.

Следует сказать, что уже созданы компьютерные технологии на основе метода семантических сетей, которые осуществляют их формирование и визуализацию. Семантическая сеть – информационная модель предметной области, имеющая вид ориентированного графа, вершины которого соответствуют объектам предметной области, а дуги (ребра) задают отношения между ними. Объектами могут быть понятия, события, свойства, процессы. Таким образом, семантическая сеть является одним из способов представления знаний. В названии соединены термины двух наук: семантика в языкознании изучает смысл единиц языка, а сеть в математике представляет собой разновидность графа – набора вершин, соединенных дугами (ребрами). В семантической сети роль вершин выполняют понятия базы знаний, а дуги (причем направленные) задают отношения между ними. Таким образом, семантическая сеть отражает семантику предметной области в виде понятий и отношений [4]. Примером отечественных разработок автоматизированного построения подобных моделей может служить система «КРИТ» компании Смартвейр (Smartware) [5], которая автоматически анализирует содержание текстовых документов, извлекает информационные сущности (например, сведения о физических и юридических лицах, географических понятиях, телефонах и адресах электронной почты) и представляет их в форме ассоциативной семантической сети. Разработчики заявляют, что используя построенные для документов семантические сети, можно выполнять различные аналитические операции: поиск неявных связей сущностей, сравнение документов на «похожесть по смыслу», рубрикация документов (автоматическая и по рубрикатору, подготовленному экспертом), построение графиков распределения тем по времени (тайм-

лайнов), построение частотных характеристик тем, вычленение оценок. В случае применения подобной технологии к задаче анализа развития научно-технической сферы формируется комбинированный «экспертно-компьютерно-информационный» подход, позволяющий решить многие задачи.

Проблема применимости информационных подходов к анализу тенденций научно-технической сферы является, на наш взгляд, междисциплинарной, находящейся на стыке экономики и информатики, и по сути затрагивает создание нового инструментария экономических исследований. Данный инструментарий базируется на следующих основаниях:

- формировании в электронной форме массивов документов (научных статей, учебных пособий, обзоров и др.), которые содержательно отражают объекты и процессы научно-технической сферы и могут быть использованы как источники для формирования структурных моделей изучаемых предметных областей и получения статистических данных по результатам поиска информации;
- наличии автоматизированных средств накопления и поиска информации для получения «информационной» статистики тенденций развития научно-технической сферы;
- разработке методик комплексного применения экономической и информационной статистики для оценок анализа тенденций развития научно-технической сферы.

Данную проблему нужно рассматривать в рамках общего процесса информатизации экономических исследований как части глобального процесса информатизации науки и общества. Начало XXI века сопровождалось качественным изменением технической и информационной базы, в результате которого мы видим появление доступной для исследователей производительной персональной компьютерной техники, быстрое развитие Интернета как глобального сетевого информационного ресурса, сетевых и «суперкомпьютерных» технологий, позволяющих объединять вычислительные мощности для решения сложных задач, а также формирование национальных и глобальных электронных информационных ресурсов. С 80-х гг. прошлого века в информатике активно развивается «семантическое» направление, общую задачу которого можно выразить как создание средств извлечения смысловой информации из информационных массивов, в том числе по запросам, возникающим в ходе анализа тенденций развития научно-технической сферы. Развитие экономической информатики и экономико-математических методов привело к созданию мощной базы обработки статистических данных и экономических расчетов. Теперь возникли условия для ее расширения за счет обработки слабоформализованных документальных информационных массивов.

Несколько слов надо сказать о том, что сама научно-техническая сфера не ограничивается только процессами теоретических исследований, прикладных научно-исследовательских работ (НИР) и созданием

опытных образцов в результате опытно-конструкторских разработок. С позиции представления о жизненном цикле систем и объектов такое ограничение было бы неправильным: «научная идея», которая в рамках научно-технической сферы получает свое решение в объекте техники, новой технологии и т. п., продолжает свое развитие и в ходе подготовки серийного производства для данного объекта, и в ходе его маркетинга, и в ходе эксплуатации (модернизации), и даже в ходе утилизации (например, утилизация отходов ядерной энергетики, которая представляет собой серьезную научную задачу, связана с получением ядерного топлива и его свойствами). В информационные системы сведения о научно-технических объектах попадают с самых различных стадий их жизненного цикла, и эти информационные системы могут, образно говоря, рассматриваться как общая плоскость проекции для всех существующих объектов НИР – предметная область научно-технического развития. При этом, если в рамках статистического учета данные о научно-технических объектах представляются в обобщенных категориях (например, «ЭВМ», «Телевизоры», «Научное оборудование» и т. п.), то в информационных системах эти данные представлены на уровне своих конкретных названий, и можно отследить не только ранее существовавшие (ретроспективные) и существующие сейчас (современные), но и перспективные научно-технические объекты-идеи в их взаимосвязи. На наш взгляд, данное обстоятельство предопределяет научную потребность в создании и развитии таких информационных средств и методов, которые позволили бы исследовать накапливаемую в информационных системах и массивах слабоформализованную информацию, прежде всего документальную.

Мы видим два пути превращения информационной системы в аналитический инструментарий для экономических исследований:

- На основе наличия в информационной системе возможностей и развитых средств статистической обработки результатов информационного поиска. Полученная информация способна существенно дополнить статистическое наблюдение, а в ряде случаев и восполнить его пробелы.
- На основе наличия в информационной системе возможностей и развитых средств для выявления семантической (смысловой) информации о структуре предметной области научно-технического развития. К таким средствам можно отнести разработанный в начале 80-х годов прошлого века В. И. Горьковой и Э. А. Бороховым «Тезаурус релятивной лексики» [6], на базе которого можно построить список из 100 глаголов русского языка, выражающих в тексте основные структурные и логические отношения между объектами. Это позволяет существенно упростить задачу анализа текстов для формирования набора объектов и матрицы взаимосвязей между ними при создании компьютерной информационной модели предметной области. Полученная информация способна дополнить, систематизировать и в ряде случаев заменить экспертные оценки.

В определенной степени можно говорить о формировании нового вида исследований – «информационно-экономических» исследований тенденций развития научно-технической сферы, которые используют различные методы информационного анализа, статистического анализа, научно-технического прогнозирования, моделирования знаний для общей цели – изучения характера развития конкретного направления научных исследований, сложного объекта или технологии. В ходе информационно-экономических исследований формируются два результата: а) новые статистические данные, получаемые на основе исследований по информационным массивам, которые дополняют данные официального статистического наблюдения, и б) данные для информационной модели, раскрывающей структуру изучаемого экономического (научно-технического) объекта. Эти два результата взаимосвязаны, поскольку дать полную оценку тенденции развития конкретного элемента информационной модели можно только с учетом его связей с другими элементами и статистической оценки динамики развития.

В экономических исследованиях научно-технической сферы информационные модели чаще всего выполняли временную описательно-иллюстративную функцию, раскрывая состав и структуру изучаемых экономических процессов, объектов или ситуаций. Если кратко остановиться на истории применения информационного моделирования в экономических исследованиях, то можно отметить несколько приметных вех. Во-первых, попытка создания автоматизированного логико-смыслового информационного фонда (ЛСИФ) для поддержки подготовки решений по проблемам управления наукой и техникой [7]. Под «автоматизированным логико-смысловым информационным фондом» понималась компьютерная система, помогающая смоделировать последствия принимаемых решений в экономике на основе совокупности представленных в ней знаний и зависимостей. Этот проект можно рассматривать как гипертекст, идеи которого один из авторов (М. М. Субботин) в 1990-х годах развил в аналитической системе CrossReader. В 70-80-е гг. прошлого века информационные модели становятся инструментом для исследования жизненных циклов организационно-экономических систем и объектов (Н. И. Комков, Г. Г. Балаян и др. [8]). В моделировании управленческих ситуаций получило развитие направление исследований с использованием логико-лингвистических моделей и специальных языков описания для решения сложных управленческих задач (Д. А. Поспелов и др. [9]). К периоду 1980-90 гг. относятся попытки применения аппарата структурного и логико-семантического моделирования для решения некоторых задач целеполагания при разработке целевых комплексных программ, в частности, предлагались подходы к формированию дерева целей на базе разработки алгоритма структурного моделирования (Д. П. Соломатин [10]) и логико-семантического моделирования структуры проблемной области целевой комплексной программы в форме семантической сети (Ю. В. Грум-Гржимайло [11]). В последнем случае для выявления связей между объектами было опробовано применение релятивной лексики, о которой говорилось выше. Это же исследование показало, что для созда-

ния моделей большой размерности нужна качественно иная программно-техническая база, в противном случае можно было бы говорить только о фрагментарном построении модели и ее экспериментальном характере. В 1990-е гг. информационное моделирование начинает использоваться для создания сложных экономических информационных систем для мониторинга экономического развития. Такие системы часто строились на основе геоинформационных систем и позволяли, например, «увидеть» экономику региона в различных разрезах, показать динамику ее развития. Однако в целом уровень технической базы того времени не позволял работы со сложными многовариантными информационными моделями, требовал их упрощения, и это стало основной причиной экспериментального характера многих работ. Сегодня техническая база позволяет решать подобные задачи, в том числе и очень сложные, на основе объединения компьютерных мощностей, создания суперкомпьютеров и т. д.

Информационно-экономические исследования, о которых мы уже говорили выше, связаны с информационным моделированием. Мы наблюдаем наличие определенного практического опыта, который позволяет говорить о формировании технологии проведения информационно-экономических исследований развития научно-технической сферы, постепенном формировании методической базы, программного обеспечения и т. д. Среди практическим наработок можно назвать:

- Исследования достижимости заданных индикаторов научно-технического развития. Примером может служить проведенное и описанное в [12] исследование российских патентных массивов и массивов отчетов по НИР и диссертациям по тематике ИКТ как технологической основы информационного общества.
- Исследования по выявлению приоритетных направлений научно-технического развития на основе статистической обработки массивов патентной информации [13].
- Исследования связей мирового научного сообщества на основе статистики социтирования по библиографическим информационным массивам [14].
- Исследования по методам моделирования структуры предметной области научно-технической сферы как основы формирования системы целей для научно-технических программ и проектов [11].

В заключение еще раз обратим внимание на то, что информационная модель может быть создана как на основе экспертных оценок, так и на основе документированного знания. Понимая, что **информационная модель «отчуждает» знания**, эксперты крайне неохотно участвуют в опросах с подобной целью, и поэтому источником знаний для формирования информационной модели должен стать документальный информационный массив, содержащий публикации и рефераты, а информационная система должна уметь извлечь и структурировать эти знания [15]. На протяжении почти тридцати лет информационные подходы постепенно занимают свою нишу в системе экономических исследований по целому ряду направлений, в том числе связанных с исследованиями

тенденций научно-технической сферы, помогая восполнить неполноту экспертных оценок и официального статистического наблюдения. Качественные изменения, происходящие в информатизации науки, формировании ее информационной инфраструктуры, не могут не влиять на методику исследований. Так, применение информационных подходов потребует наличия доступа к информации, информационной техники и специальных информационно-поисковых систем. С одной стороны, это приведет к возрастанию накладных расходов, с другой – сократит сроки проведения исследований и повысит их качество. Можно ли ожидать, что применение информационных походов в определенной степени изменит сложившуюся парадигму в области экономического анализа в сторону формирования междисциплинарного инструментария, характерного для науки в условиях информационного общества?

Источники и литература

1. Большой психологический словарь / Сост.: Мещеряков В., Зинченко В. М.: Олма-пресс, 2004.
2. *Глушков В. М.* Гносеологическая природа информационного моделирования // Вопросы философии. 1968. № 10. С. 13–18.
3. ИСО 10303-1:1994, ст. 3.2.21.
4. http://ru.wikipedia.org/wiki/Семантическая_сеть
5. http://www.smartware.ru/semnet.html
6. *Горькова В. И., Борохов Э. А.* Реферат в системе научной коммуникации. Направления совершенствования лингвистических и структурных характеристик. М.: ВИНИТИ, 1987 (Итоги науки и техники. Сер. Информатика / ВИНИТИ. Т. 11).
7. Возможности использования автоматизированного логико-смыслового информационного фонда при подготовке решений по проблемам управления наукой и техникой. М.: ВНИИПОУ, 1982. – 46 с.
8. *Балаян Г. Г., Жарикова Г. Г., Комков Н. И.* Информационно-логические модели научных исследований. М.: Наука, 1978; *Балаян Г. Г.* Информационное моделирование научно-технических программ. М.: Наука, 1987; *Комков Н. И.* Модели программно-целевого управления. М.: Наука, 1981.
9. *Поспелов Д. А.* Логико-лингвистические модели в системах управления. М.: Энергоиздат, 1981.
10. *Соломатин Д. П.* Математическое и программное обеспечение диалоговой системы структурного моделирования. М.: ВНИИСИ, 1982 (Препринт).
11. Информационная технология формирования целей научно-технических программ на основе логико-семантического моделирования // Новые формы связи науки с производством. М.: Наука, 1992; *Грум-Гржимайло Ю. В.* Экспертно-информационная технология анализа и подготовки стратегических решений: Дисс. ... канд. эконом. наук. М.: РЭА им. Г. В.Плеханова, 1997.

12. *Попов С. В., Грум-Гржимайло Ю. В., Сергеева В. В.* Оценка достижимости контрольных значений показателей, заданных в «Стратегии развития информационного общества в Российской Федерации», на основе библиометрического анализа // Наука. Инновации. Образование. Вып.8. М.: Языки славянской культуры; РИЭПП, 2009. См. наст. сб.
13. *Попов С. В.* Корпоративный форсайт и конкурентная разведка // Наука. Инновации. Образование. Вып. 6. М.: Языки славянской культуры, 2008.
14. *Шапошник С. Б.* Международное научное сотрудничество России: библиометрическое исследование // Науковедение. 1999. № 1. С. 157–171.
15. *Грум-Гржимайло Ю. В., Попов С. В., Поляков М. Н.* Информационный мониторинг структуры проблемной области для поддержки принятия управленческих решений в научно-технической сфере // Наука. Инновации. Образование. Вып. 6. М.: Языки славянской культуры; 2008. С. 223–236.